Geophysical Research Abstracts, Vol. 9, 02487, 2007
SRef-ID: 1607-7962/gra/EGU2007-A-02487

Pruning of hidden nodes in support vector networks on flood forecasting

S.-T. Chen and P.-S. Yu

Department of Hydraulic and Ocean Engineering, National Cheng Kung University, Taiwan (yups@ncku.edu.tw / Fax: +886 6-2741463)

The kind of support vector machine (SVM) that uses the radial basis function (RBF) as a kernel is structurally the same as an RBF neural network. The support vectors (SVs) in SVM can be regarded as the centers of the hidden nodes in an artificial neural network (ANN), and the number of SVs is the number of hidden nodes. Pruning of neural networks can increase the speed of calculation, reduce the storage requirements and enhance generalization ability. The pruning of ANNs and its application have been widely discussed. However, the pruning of support vector networks is relatively few reported.

This study proposed an approach to the pruning of SVs in SVM networks according to the parameter characteristics. The error tolerance ε in the support vector regression model has very high correlation with the number of SVs, and the correlation is quite consistent, regardless of the values of other parameters (the penalty constant C and the RBF parameter γ). This property of SVM can be used to construct the support vector networks with desired number of SVs according to an assigned ε value. The advantage of this method is that the resulting support vector network still has its optimal network structure that is mathematically derived. The SVs are automatically reduced by the SVM algorithm simultaneously, but not to prune afterwards the SVs in the already derived optimal network.

The proposed pruning method was applied to a case study of real-time flood stage forecasting in the Lan-Yang River, Taiwan. Five pruned SVM models with respect to different percentages of the number of SVs were built. Results of these models to forecast multiple-hour-ahead flood stages were presented, and subjects relating to the pruning of SVM networks (such as the identification of support vectors and the

parameter sensitivity) were discussed.